

EP 0 518 443 B1 (11)

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent: 18.05.2005 Bulletin 2005/20

(51) Int Cl.7: C12N 15/10

(21) Application number: 92201711.6

(22) Date of filing: 30.10.1985

(54) Recombinant proteins of viruses associated with lymphadenopathy syndrome and/or acquired immune deficiency syndrome

Rekombinante Proteine von mit Lymphadenopathiesyndrom und/oder erworbenem Immunschwächesyndrom assoziierten Viren

Protéines recombinantes de virus associé au syndrome lymhadenopatique et/ou syndrome d'immunodéficience acquise

(84) Designated Contracting States: AT BE CH DE FR GB IT LI LU NL SE

(30) Priority: 31.10.1984 US 667501 30.01.1985 US 696534 06.09.1985 US 773447

- (43) Date of publication of application: 16.12.1992 Bulletin 1992/51
- (60) Divisional application: 02077411.3 / 1 245 678
- (62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC: 85307860.8 / 0 181 150
- (73) Proprietor: CHIRON CORPORATION Emeryville, California 94608 (US)
- (72) Inventors:
 - · Luciw, Paul A. Emeryville, California 94608 (US)
 - Dina, Dino Oakland, CA 94602 (US)
 - Steimer, Kathelyn Benicia, California 94510 (US)

- · Pescador, Ray Sanchez Oakland, California 94605 (US)
- · George-Nascimento, Carlos Danville, California 94526 (US)
- · Parkes, Deborah No. 104 Oakland, California 94610 (US)
- Hallewell, Rob San Francisco, California 94109 (US)
- Barr, Philip J. Orinda, California 94563 (US)
- Truett, Martha Oakland, California 94611 (US)
- (74) Representative: Hallybone, Huw George Carpmaels and Ransford, 43 Bloomsbury Square London WC1A 2RA (GB)
- (56) References cited:

EP-A-0 139 216

EP-A- 0 152 030

EP-A-0 173 529

EP-A- 0 185 444

 SCIENCE, vol. 225, no. 4664, August 24, 1984 Washington DC J.A. LEVY et al. "Isolation of Lymphocytopathic Retro- viruses from San Francisco Patients with AIDS", pages 840-842

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

25

Technical Field

[0001] This invention is in the field of lymphadenopathy syndrome and/or acquired immune deficiency syndrome.

Background Art

[0002] With the discovery of human T-cell lymphotropic virus-I (HTLV-I) as an infectious agent in humans, it was established that retroviruses could infect humans and could be the etiological agent of disease. After HT-LV-I was established, a second retrovirus of the same family, HTLV-II was found in a hairy cell leukemia established strain. Since that time, other human retroviruses have been isolated which are associated with lymphadenopathy syndrome (LAS) and/or acquired immune deficiency syndrome (AIDS) victims. Various retroviruses have been isolated from individuals with AIDS (sometimes called HTLV-III) or LAS (sometimes called LAV). See for example, Barre-Sinoussi, et al, Science (1983) 220:868-871 and Montagnier, et al, Cold Spring Harbor Symposium (1984) in press; Vilmer, et al, Lancet (1984) 1:753, Popovic, et al, Science (1984) 224:497 and Gallo, et al, Science (1984) 224:500. A comparison of HT-LV-III and LAV may be found in Feorino, et al, (1984), supra. See also, Klatzman, et al, Science (1984) 225: 59-62, Montagnier, et al, ibid (1984) 63-66, and the references cited therein for a survey of the field. A general discussion of the T-cell leukemia viruses may be found in Marx, Science (1984) 224:475-477. Levy, et al, Science (1984) 225:840-842 report the isolation of ARV (AIDS-associated retroviruses).

[0003] For the purposes of this application, these viruses (HTLV-III, LAV, and ARV) will be generically referred to as human T-cell lymphotropic retrovirus (hTLR). The hTLRs may be shown to be of the same class by being similar in their morphology, serology, reverse transcriptase optima and cytopathology, as identified in the above references. For example, the reverse transcriptase prefers Mg⁺², and has a pH optima of about 7.8.

Disclosure of the Invention

[0004] The invention is defined in the claims.

Brief Description of the Drawings

[0005]

Figure 1 is a restriction map of proviral DNA (ARV-2).

Figure 2 is the nucleotide sequence of ARV-2(9B). The amino acid sequences for the products of the gag, pol, and env genes are indicated. The U3, R,

and U5 regions of the LTRs are also designated. The cap site is position +1. A 3 bp inverted repeat at the ends of the LTR, the TATA box at position -29, the sequence complementary to the 3'-end of the tRNAlys at position 183, and the polyadenylation signal at position 9174 are underlined. The overlines indicate the amino sequences determined from virion proteins. The nucleotides at the beginning of each line are numbered, and the amino acids at the end of each line are indicated.

Modes for Carrying Out the Invention

[0006] The hTLR DNA sequences are synthesized at least in part and may be used for expression of polypeptides which may be a precursor protein subject to further manipulation by cleavage, or a complete mature protein or fragment thereof. The smallest sequence of interest, so as to provide a sequence encoding an amino acid sequence capable of specific binding to a receptor, e.g., an immunoglobulin. will be 21 bp, usually at least 45 bp, exclusive of the initiation codon. The sequence may code for any greater portion of or the complete polypeptide. or may include flanking regions of a precursor polypeptide, so as to include portions of sequences or entire sequences coding for two or more different mature polypeptides. The sequence will usually be less than about 5 kbp, more usually less than about 3 kbp. [0007] Sequences of particular interest having open reading frames (Figure 5) define the structural genes for the gag proteins (p16 and p25) and the env protein. It is to be understood that the above sequences may be spliced to other sequences present in the retrovirus, so that the 5'-end of the sequence may not code for the Nterminal amino acid of the expression product. The splice site may be at the 5'-terminus of the open reading frame or internal to the open reading frame. The initiation codon for the protein may not be the first codon for methionine, but may be the second or third methionine, so that employing the entire sequence indicated above may result in an extended protein. However, for the gag and env genes there will be proteolytic processing in mammalian cells, which processing may include the removal of extra amino acids.

[0008] In isolating the different domains the provirus may be digested with restriction endonucleases, the fragments electrophoresed and fragments having the proper size and duplexing with a probe, when available, are isolated, cloned in a cloning vector, and excised from the vector. The fragments may then be manipulated for expression. Superfluous nucleotides may be removed from one or both termini using Ba131 digestion. By restriction mapping convenient restriction sites may be located external or internal to the coding region. Primer repair or in vitro mutagenesis may be employed for defining a terminus, for insertions, deletion, point or multiple mutations, or the like, where codons may be changed, either cryptic or changing the amino acid. re-

20

40

striction sites introduced or removed, or the like. where the gene has been truncated, the lost nucleotides may be replaced using an adaptor. Adaptors are particularly useful for joining coding regions to ensure the proper reading frame.

[0009] The env domain of the hTLR genome can be obtained by digestion of the provirus with <u>Eco</u>RI and <u>KpnI</u> and purification of a 3300 base pair (bp) fragment, which fragment contains about 400 bp of 5' non-coding and about 200 bp of 3' non-coding region. Three different methionines coded for by the sequence in the 5' end of the open reading frame may serve as translational initiation sites.

[0010] Digestion of proviral sequences with <u>SacI</u> and <u>EcoRV</u> provides a fragment of about 2300 bp which contains the gag domain and a second small open reading frame towards the 3' end of the gag region. The gag domain is about 1500 bp and codes for a large precursor protein which is processed to yield proteins of about 25.000 (p25), 16.000 (p16) and 12,000 (p12) daltons. Digestion with <u>SacI</u> and <u>BgIII</u> may also be used to obtain exclusively the gag domain with p12, p25 and partial p16 regions.

[0011] The hTLR DNA sequences may be labeled with isotopic or non-isotopic labels or markers and be used as DNA probes to detect the presence of native hTLR nucleotide sequences in samples suspected of containing same.

[0012] The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

1. AIDS related virus-2 (ARV-2) purification and preparation of viral RNA.

[0013] HUT-78 cells infected with ARV-2 (ATCC Accession No. CRL B597, deposited on August 7, 1984) were obtained from Dr. Jay Levy, University of California, San Francisco. Cultures were grown for two weeks in RPM1 medium with 10% fetal calf serum. Cultures were centrifuged at 2 Krpm for 1 hr at 4°C using a SW-2B rotor. The pellet, containing the virus, was resuspended in 10 mM Tris-HCI, pH 7.5 on ice. The resuspended pellet was treated with 10 µg of DNase (Boehringer Mannhein) and was layered onto a linear sucrose gradient (15-50% in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 20 mM NaCl). The gradient was spun at 34 Krpm for 4 hr at 4°C, in SW-41 rotor. Five 2.5 ml fractions were collected and an aliquot of each was electrophoresed in a 1% agarose, 5 mM methyl mercury hydroxide gel (Bailey and Davidson. Anal Biochem (1976) 70:75-85) to determine which contained the 9 kb viral RNA. The fraction containing the viral RNA was diluted: to 10 ml in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA and was centrifuged at 34 Krpm for 2 hr at 4°C. The pellet was resuspended in 20 mM Tris-HCl, pH 7.6, 10 mM EDTA, 0.1% SDS, and 200 µg/ml proteinase K. Incubation was carried out for 15 min at room temperature. The mixture was extracted with phenol and the aqueous phase was made 400 mM NaCl and precipitated with ethanol. The pellet was resuspended in water and stored at -70°C.

[0014] To purify the viral RNA from the nucleic acid pellet obtained as described above, a sample was electrophoresed in a low-melting 1% agarose gel containing 5 mM Methyl mercury hydroxide. After electrophoresis, the gel was stained with 0.1% ethidium bromide and nucleic acid bands were visualized under UV light. The region corresponding to 9 kb was cut from the gel and the agarose was melted at 70°C for 2 to 3 min in three volumes of 0.3 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA. The mixture was extracted with an equal volume of phenol. The aqueous phase was reextracted with phenol and was precipitated with ethanol. The pellet was washed with cold 95% ethanol, air dried, resuspended in water and stored at -70°C until use. One hundred ml of culture medium yielded 0.5 to 1 µg of purified RNA.

2. Synthesis of labeled homologous viral probe.

[0015] A 32P-labeled cDNA was made to the gel purified viral RNA using random primers (calf thymus primers) prepared as described in Maniatis, et al, A Laboratory Manual, Cold Spring Harbor, NY, 1982. The reaction mixture contained 2 µl of 0.5 M MgCl₂; 5 µl of 0.1 M dithiothreitol; 2.5 µl each of 10 mM dATP, 10 mM dGTP and 10 mM dTTP; 2.5 µl calf thymus primer (100A₂₆₀/ ml); 0.5 μg viral RNA: 5 μl of actinomycin D (200 μg/ml); 10 μl of ³²P-dCTP (> 3000 Ci/mmole, 1 mCi/ml) and 1 μl of AMV reverse transcriptase (17 units/μl) in a 50 μl reaction volume. The reaction was incubated for 1 hr at 37°C. The probe was purified away from free nucleotides by gel filtration using a Sephadex G50 column. The void volume was pooled, NaCl was added to a final concentration of 400 mM and carrier single-stranded DNA to 100 µg/ml, and the cDNA was precipitated with ethanol. The pellet was resuspended in water and incorporated 32P counts were determined.

3. <u>Detection of ARV sequences in polyA+ RNA prepared</u> from infected HUT-78 cells.

[0016] PolyA+ RNA was prepared from HUT-78 cells infected with ARV-2, ARV-3 or ARV-4 (three different isolates from three different AIDS patients) and from uninfected HUT-78 cells. The polyA, RNA was electrophoresed on 1% agarose gels containing 5 mM methyl mercury hydroxide (Bailey and Davidson, supra), was transferred to nitrocellulose filters, and hybridized with the homologous probe prepared as described in Section 2. Hybridizations were carried out in 50% formamide, 3 x SSC at 42°C. Washes were at 50°C in 0.2 x SSC. A 9 kbp band was present in all three samples of infected HUT-7B cells. This band was absent in polyA+ from uninfected cells.

4. <u>Detection of ARV sequences in infected and non-infected HUT-78 cells.</u>

[0017] High molecular weight DNA (chromosomal) was prepared from cultures of HUT-78 cells infected with ARV-2 and from non-infected HUT-78 cells following the procedure of Luciw, et al, Molec and Cell Biol (1984) 4:1260-1269. The DNA was digested with restriction enzyme(s), electrophoresed in 1% agarose gels and blotted onto nitrocellulose following the procedure described by Southern, (1975), supra. Blots were hybridized with the ³²P-labeled probe (10⁶ cpm/blot) in a mixture containing 50% formamide, 3 x SSC, 10 mM Hepes, pH 7.0, 100 μg/ml denatured carrier DNA, 100 μg/ml yeast RNA and 1 x Denhardt's for 36 hr at 42°C. Filters were washed once at room temperature in 2 x SSC and twice at 42°C in 0.2 x SSC, 0.1% SDS. Filters were air dried and exposed to X-Omat film using an intensifying screen.

[0018] The homologous ³²P-probe to ARV-2 hybridized specifically to two bands in the DNA from infected cells restricted with <u>SacI</u>. These bands were absent when DNA of non-infected cells was used, indicating that the probe is hybridizing specifically to infected cells presumably to the provirus integrated in the chromosomal DNA. The molecular weight of the bands is approximately 5 kb and 3 kb.

[0019] In order to determine if different enzymes would cut the proviral sequence, several other restriction digestions of the cell DNA were carried out using EcoRI, Sphl or Kpnl or double digestions using two of them. Southern results show specific bands hybridizing when DNA of infected cells is used. Figure 1 shows a schematic map of the positions of restriction enzyme sites in the proviral sequence, and indicates fragment sites.

5. Cloning of proviral ARV-2 DNA.

[0020] High molecular weight cell DNA from infected HUT-78 cells was prepared following the procedure of Luciw, et al, supra. The DNA was digested with EcoRI, which cuts once in the previrus, centrifuged in a sucrose gradient and fractions corresponding to 8-15 kb were pooled, dialyzed and concentrated by ethanol precipitation. The bacteriophage λ derivative cloning vector, EM-BL-4 (Karn, et al, Methods Enzymol (1983) 101:3-19) was digested to completion with a mixture of EcoRI, BamHI and Sall restriction enzymes and the DNA then deproteinized by phenol-chloroform extraction, precipitated with cold ethanol and resuspended in ligation buffer. The EMBL-4 phage DNA and EcoRI digest of cellular DNA were mixed and ligated and the resultant recombinant phage genomes packaged in vitro. After phage infection of λ -sensitive E. coli (DP50supF), about 500,000 phage plaques were transferred onto nitrocellulose filters, DNA was fixed and the filters were screened with a homologous 32P-probe prepared as described in Section 2. Eleven recombinant phage out of 500,000 phage annealed in the initial double-lift screening method (Maniatis, et al, Molecular Cloning, A Laboratory Manual, NY, 1982) to viral cDNA probe, and these were further plaque-purified and propagated in large liquid cultures for preparation of recombinant DNA. Plaque-purified phage containing ARV DNA were propagated in liquid culture in E. coli DP50supF; phage particles were harvested and banded in CsCl gradients and recombinant phage DNA was prepared by phenol extraction followed by ethanol precipitation (Maniatis, et al, supra). One μg of purified phage DNA was digested with restriction enzymes, electrophoresed on 1% agarose gels, and visualized with ethidium bromide under ultraviolet light. The DNA from these gels was transferred to nitrocellulose and annealed with viral cDNA probe.

[0021] One of the 11 phage, designated λ ARV-2(9B). was deposited at the ATCC on 25 January 1985 and given Accession No. 40158. λ ARV-2(9B) contained an insertion of full-length proviral DNA along with flanking cell sequences. Digestion of λ ARV-2(9B) DNA with Sacl yielded viral DNA fragments of 3.8 kb and 5.7 kb. EcoRl digestion of λ ARV-2(9B) produced virus containing DNA species at 6.4 kb and 8.0 kb; a double digest of Sacl and EcoRl gave viral DNA fragments at 3.8 kb and 5.4 kb. This pattern is consistent with that of a provirus linked to cell DNA.

[0022] In addition to λ ARV-2(9B), phage was obtained that (1) possessed the left half of the viral genome from the EcoRI site in viral DNA extending into flanking cell DNA (λ ARV-2(8A)) and (2) phage that had the right half of the viral genome (λ ARV-2(7D)) from the EcoRI site in viral DNA extending into flanking cell DNA. Bacteriophages λ ARV-2(7D) (right) and λ ARV-2(8A) (left) were deposited at the ATCC on October 26, 1984 and given Accession Nos. 40143 and 40144, respectively.

Polymorphism.

[0023] To measure the relatedness of independent ARV isolates, restriction enzyme digests of DNA from HUT-78 cells infected with ARV-3 and ARV-4 were analyzed with the probe made from cloned ARV-2 DNA. The SacI digest of ARV-3 DNA was similar to that of ARV-2 whereas the HindIII digests displayed different patterns. The SacI digest and the PstI digest of ARV-4 DNA differed from the corresponding digests of ARV-2 DNA. The intensity of the annealing signals obtained with ARV-3 and ARV-4 samples was much lower (about 10-fold less) than that for ARV-2 DNA probably as a result of the fact that fewer cells were infected in the ARV-3 and ARV-4 cultures. The viral-specific DNA fragments produced by Sac1 treatment of ARV-3 and ARV-4 DNA totaled 9.0-9.5 kbp, a value similar to that of ARV-2 and in consonance with the RNA genome sizes.

20

25

30

35

40

7. Sequencing of proviral DNA.

[0024] Fragments or subfragments of ARV-2 DNA from λ phage 9B were prepared and cloned into M13 according to conventional procedures (Maniatis, et al, supra). Sequencing was performed according to Sanger, et al, <u>Proc Natl Acad Sci USA</u> (1977) 74:5463, using the universal M13 primer or chemically synthesized primers complementary to ARV-2 sequence. The sequence is shown in Figure 2.

Claims

- A process for the preparation of an isolated DNA polynucleotide comprising a fragment of at least 21 bp from the gag or env region of the ARV-2 sequence of Figure 2, wherein the DNA polynucleotide is chemically synthesised at least in part.
- The process of claim 1, wherein the DNA polynucleotide comprises a fragment of at least 21 bp from the gag region of the ARV-2 sequence of Figure 2.
- 3. The process of claim 1, wherein the DNA polynucleotide comprises a fragment of at least 21 bp from the env region of the ARV-2 sequence of Figure 2.
- The process of any preceding claim, wherein the DNA polynucleotide is labelled.
- The process of claim 4, wherein the DNA polynucleotide is labelled with an isotopic label.
- The process of claim 4, wherein the DNA polynucleotide is labelled with a non-isotopic label.

Patentansprüche

- Verfahren zur Herstellung eines isolierten DNA-Polynucleotids, das ein Fragment von mindestens 21 bp aus der gag-Region oder der env-Region der ARV-2-Sequenz gemäß Figur 2 umfaßt, wobei das DNA-Polynucleotid mindestens teilweise chemisch synthetisiert wird.
- Verfahren nach Anspruch 1, wobei das DNA-Polynucleotid ein Fragment von mindestens 21 bp aus der gag-Region der ARV-2-Sequenz gemäß Figur 2 umfaßt.
- Verfahrennach Anspruch 1, wobei das DNA-Polynucleotid ein Fragment von mindestens 21 bp aus der env-Region der ARV-2-Sequenz gemäß Figur 2 umfaßt.
- 4. Verfahren nach einem der vorhergehenden Ansprü-

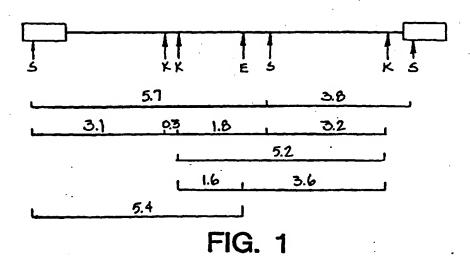
che, wobei das DNA-Polynucleotid markiert ist.

- Verfahren nach Anspruch 4, wobei das DNA-Polynucleotid mit einem isotopischen Marker markiert ist.
- Verfahren nach Anspruch 4, wobei das DNA-Polynucleotid mit einem nicht-isotopischen Marker markiert ist.

Revendications

- Procédé pour la préparation d'un polynucléotide d'ADN isolé comprenant un fragment d'au moins 21 pb de la région gag ou env de la séquence ARV-2 de la figure 2, dans lequel le polynucléotide d'ADN est chimiquement synthétisé au moins en partie.
- Procédé selon la revendication 1, dans lequel le polynucléotide d'ADN comprend un fragment d'au moins 21 pb de la région gag de la séquence ARV-2 de la figure 2.
- Procédé selon la revendication 1, dans lequel le polynucléotide d'ADN comprend un fragment d'au moins 21 pb de la région env de la séquence ARV-2 de la figure 2.
- Procédé selon l'une quelconque des revendications précédentes, dans lequel le polynucléotide d'ADN est marqué.
- Procédé selon la revendication 4, dans lequel le polynucléotide d'ADN est marqué avec un marqueur isotopique.
- Procédé selon la revendication 4, dans lequel le polynucléotide d'ADN est marqué avec un marqueur non isotopique.

5

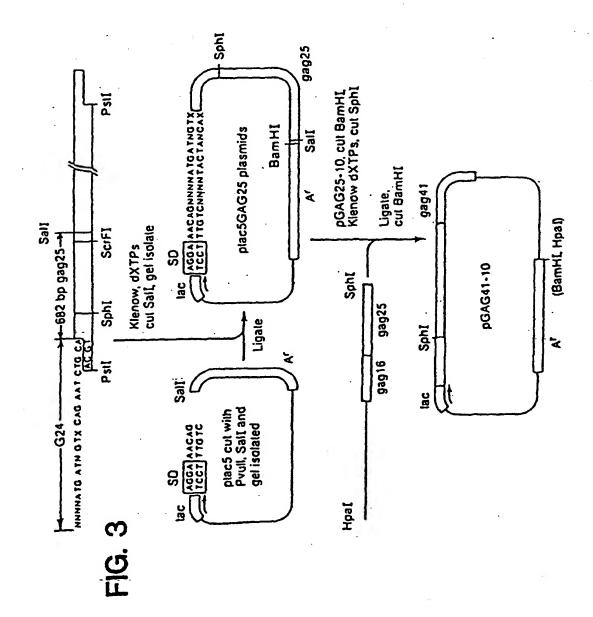


```
-413 CLECHWERCLAMALINERICCEWWRITERCOTOPICATICTIC CONTROL LANGE CONTROL LA CONT
                           33974334637773843677738436777367773677738788878782782782782782879386778677867786778677867778677787
                 BACCCITYTAGTCABTBTBEADAAATCTTETABCAB <u>TBECCECCCAACGAGTGA</u>CCECAAAGCCAAGATCCAGAACCCAGGACCTCCTCTCGAGGACTCGCTTGCTGAAGCCCCACAG
               Lys Stadys Stady Advantage Stady Sta
                Argelailalooliygialootlaprolorlaabiylortigeliysortigelooliyasatbiyalalatbilootiyreysvalkiselabiyliaasptaliysaspthi 101
Araxamatatteegaxaeetakaeeesiceettaabaxaeeatkaemaeattaehekattatathakaetaeeatkeettetattetetakaetamaeeatkeetetaka
                 Lystiväl aloogiolys 11 eti sõistiosi odjolys sorlyslyslysäisti oti sälsäisäisäi sälsti yöptiyäroser sorti sersi sersisäistyy proii otil
Aaceaacettararaa taraceaarar aanaraanetaisia aanararaksaketaketekäiset oli yöptiyäroser sartiasitareetaattareet
                Ginasmi ducioci y cinas tal discòna do lla serproar q'arlevasad al prolity eta i tal duciul y cha per proci dual i lepromotada serà fala u
Cacaace facacecasa del facate acceca farese e tacaace i faca de cala desta da cacaace e terme de la cacacect de
               Sarciuslyalatheprosloasploussatherotlouasatherologyslykistiaalpalokettikanstloulystiethejisasotloulualaalatiotepaspaegust
Teakaateaeeeaeeeeaeastiiaaaeaeeateetaaaeaeagtosseeaeateteaeeaeatetataaaaaaeaatateaaeaaeeteeaeaateegatabate
                llafratalBiyGlullaTyrlysärgTrpllallalagiyleukialysilatalärgmatTyrBerfrotbrBerltalaukipilaärgülaElyfrolysGlufrofbeärgäib. 30%
ATCCCABTAGAAAATTTATAAAAATEGATAAYECTGBATTAAATAAAATABTAABAATTÄTAEECCTAECAGCATTETGGACATAAGAAAGAACCATTAABGAACCCTTTABBGA
                 Tyrtolaspärgpoolyriysihriqqärasioglaalasorolaasptoliysässippästingoloteloglogtolasaaloasopraaspiysiysiörilologiyo
Tatttaaregettetataaateetaigaaregaaretteleareatetaaaaattearegaretaetetetestikkingologitettääteetaittetaitetaite
                Al alous ly proal oal other local del uno ther the alocy slogly voled y ply prosly mist, year of a fight on the local of the approal oa a general of the control of the approal oa a general oa approal oa ap
                Angrolesta Tarti estelielielielieliese pisiosi eleolomas pripisiyai espasotrat leugiusi motaselouprosi ylystrolysprolysmet (i e. 103
GACCCCTCGTCACASTAASSATASSESSECAACTAASSAASCTCTATTASATACASBASCASATSATACAGTATTASAASAAATTTSCCASSAAATTOCCASSAAACTAASSAASCAAATSATAS
               61761711a6176179ballalyatalarg61atypasp61a11aprota161a11a6ys6178talyaala11a617taptal6ayta617protappota1a11a11a617ag
- aasia1118aase17ffateaaataasaacaatacaataaatact61abaatct818aastaaastalaastalaastalatastactacaastatataataataata
                 isalaulootbrēlai isētyčystērlomās pāropro ilošopāra (labiotepts) prato ilps Lomlys procijata applyprolys ta ilps Biotopolautarcia. 183
Atciett bactcasaticos tietacti tarticoccati astectati sakates ackas takactas sakatesa izsēcemas pirak paitak pa
               Sialysiislysaialeuvaisisiiseegastapsiskussislulyssiustylysiiseerkysiisetypossiaaaappotypäsatäpppopaisitalyslyslysäspeep 223
.AAAANTAAASSERTTASTASASSITTYTEETTYTTÄTTYSTENAARSERAAATTTEESTETTÄÄNTYTESTETTYTEETTYTTETTYTÄNAAASSERAA
                 folglyse pasplossic i estysiasi sapytoplys i estusiosapysissi slowlopapytopsi prodofropapplys i estalyestoppopolob (b)
Tarrattyrrattaratapprostartatarataratararraktararrattsi sapstoccattyreter i estalataratapparataraktyretytytty
               Tromots) ptyrsi ulaumi aprokaplya trothe vai si aproi i omotlou proši klyakapšou to the Valkamkopi i oši klyalouvai si plyaloukam troki palaker. 423
Relygesty kteret ta trak trak traker straker ta tratsetecker karaker kara to tekat si karaker karaker karaker k
                giojiatypatagytolystolystolegocyslyslarionapstytethitätältetheristattolegototototologatagiaelogiologislogislag
Apattiatericcaattaaattaaerastatetaaartettireaggaartaaarkaragaastaatackataarkaragaarkaragaartagaartekaaaaakkara
                 ilaloulyagiuprovainisgistaityptypäadpprosorlynäadlostaiälagivilogialyagiogiygiatyptaptypgiailolypgiagiuppopholyaasa
Tyrtaalagaarcastarattaatatattattareksetataaaarattagtartagaataragaarragograasorratyptatatatraattyptraagaarratya
                 leelys Torgiykys Tyraienym targgiyai omis Toraseaspfol Lysgieloo Torgizal tabilys fol Bor Torgiusor Itorosiyl 1967-1961 yl 1967-1985
Tellale ascalat 1982 aleel teleseet ee er aptaatgatet alakkastt aleesteet ee alakkastateleetaat at alta taleel
               PholyslouProlloGlalysQuaterfopGlahlatopfopMetGlutyrfopGlahlatbetroplupProGlutoButalAsaTheProProleutalLyslouTopfyrGlu 303
THAMETACCATACAAAEEAAACATBACAAECATGGTGGAEGAETATBACAACTACCTGAATGCGAGTTTGCAATTACCCCTCCCTTAGTGAAATTA1GGTACCAGT
                lorlielyr Birpo i i oto i birto topto typta i kipî i ali aliranggi i toplogi yiyal bi i tiptata i toplogi blysta i ali se 627
Tarana kareca tartarak karaket tepatetari erreketak tarak karaketarak tarakarak karaka tartak karaketarak karak
              CipproApplys Ser Bio Sór Fiolowal Sor Gio i le ilo Gio Cio Chen La Capaca de Capaca de
               roprove i volatel pegi oli oto tata agorcyba splysčysčiači poli valobeni sči poli oto taja spčys sopročiy i to trpči elova 192
Cacci otagi se cali og pata stag cacci oci oci sa sala sterecta sakce stereca stag se specio stag stereca stag t
```

FIG. 2-1

```
ipāri prāspēroloutrojyssī pērodislyslouloutrojystystustototoi pietimāspāsa Sorāspidelysta trotprodrylysā algyblo 110—983
Acarcamaticecettisgas argecegemas ettetetemas prometas taptas taptas taptas etamas taticecemas kallandemas kal
                  <sup>a</sup>aptyrétylylétamogát aétyáspálpépetelátásberárgétmáspét<mark>uáspán</mark>
Sattatesanakakatesekaseteateatteteteselagyásakassatangsattagakatesanaketttagtanaka<del>le</del>satátgtatattt<u>kaakaakaetaa</u>gg
        TTTTATABAÇATCACTATGAAAGTACTCATCCAAGAGTAAGTTCAGGAAGTACACATCCCCCTAGGGAATGCTAAATTGGTAATAACAACATATTGGGGCTCTGCATACAGAAAA
        49.00.000 to 19.00.000 to 19.00.000 to 19.00.000 to 19.000 to 19.0
            taararakanectettaracattitoctabbeeratereteratarectirrereratatatatetatatereratat<del>aererataetaerereratrererakaetaetae</del>
       AACTECTETTETTEATITEATATEAAATTEESTETTEAACATASEAAAATAEGEATTATTEAACAAGAAATEAAGAATEAACTEAGEETAAGTAGATETAACTAGTETAA
              LyselyThrargargasaTyrelamislooTrakrgTraelyThraolooloolooglyMotloomeelleCysSorAlaThreUnlyslooTratoThrTolTyrTyre)yYalProtal 51
AAGGGGACCACEAGGAATTATCAGCACTIGTGGAGATGGGGCACCTTGCTGCTGGGATGTTGATEATCTGTAGTGCTACAGAAAAATTGTGGGTGACACTTTATTATAGAGAAAATTGTGGGT
       trplyiblualatartartartarloopaetyiälaseraipalaargalatyraaptbebluvalulaabatattrpalatbeulsalacysvolprotoraipproajaproglobluval 91
Tesaaasaagtactaccactetattiteteeateabatgetasaegatatbatacaaagetagtaattattegeegeacacatecteeteteteteegeegeabaccacag
    YallousiyasavaltarsidaasaParaamottrpijsäsääsamotvalsiusiamatsiasidainispilajioSorioutopäagsinSorioulysProtysvaltystauthePro 231
S7477664aarf575428aaaar777744547578aaAaa7445457454454545454545454545454777485047674455476744654774455477445547
    11 of applicable cyclopic years 1 september of population in the process of the process of the second secon
    ASPEINSERSERGI VELVAS PROGIUTIOTATROLUTS SORPHAANCYSÄRGEI VETUPAPPAD PROTYTATALALUTAALLE OUPHAANAANIN TYVÄRGLOVÄNNUS SIDRELU ALL
AATCAATEETEAGGAGGGAGECCAGAAATTETAATEENEAGAGTITTAATTETAGGGGGAATTITTETAETGTAATACAALAGAGETTTAATALAILAENETTAATAEATE
    Typlystal Helys (1 of 1 oproleug) y 11 oli oprotoplyski slyskegårgteltel flokegålelysåpgåletsigt y 11 otal 17 y 11 otal 17
    ai aligiysgy tarmetgiyai atai Seyleg Tarleg Tartai Giaal aapggi alegleg Seysi atai Giagiaalaadaa kaleglegigai ai iggigai agiag iaga bay
Geografiaag kataggag tataggggas 18 teleg getaggg taregggearaa tartaty 67 tegastaaraa taragagaa tatag tagagtay t
    laglagialagīpyaifradipadiyilalyabialagialiakatālikalbiakatīptyrtagaapbiadialaglagiyilatabiptyesartiylylagitadya
Birtteenatelalatītaggeentengelacteenegenagateetegeentaganakatnetinaggatenklagtettabeatiteetetelalalatetattee
    Locions) of 10 lorg) abagiagiagiagi biy 18 lagi of 16 Gine olomi biyatral 18 of 17 gine trapagarti of 17 ko 17 pi 10 ily 11 o 69)
Tile tralegla teegracearla blala baaraha kataat kataat keataat teeataas tile teegat teegat tagaal teega tagaat t
           ie ilematieta iši yši ylemalityjemarą (letelpman) atalem Serti otathimarymalaryjenia jytyršerprolem Serphogia tararji oppratelpro
Icarateatan tannanna trontantitaacan tagrifficci (intertatan tannangunian nangunian taken kalentar taran terang
    LoubuserTyrkrakratomkrakalouvelouliakiakiakrathryaibivliakoutyyn skrabiytrakialoukysTyrtrykrakratousiatyrtryti
Cicticaectaecutekratokalottaetettaatireaecassakiatousaktioiseessaksaksaksessaksaksettytääktätisekratoksaktatie
   Elaciuloulyska a Serál a tal Sorteplauda a diatará a tial far Ciar ty tarán par que la facila de la degal a tyra gal a lleun la light a sa la contra de la contra del la contra de la contra del la contra del la contra de la contra del la contra de la contra del la 
           gargirargiratylogbirkyglorloglorog
Malantirencessetteseannssettitsetataratsostsocalastsetenetastatssetssetssetssetstessetataasssaaslatgaratrascass
  CENALAGE (COCT TT CCARTCAGACC TCAGGTACC TTTAMENCAN TEACTTACAGGE ACCTTAGATATTAGCCACT TTTTAMAGAGAAGGGGGGA CTGGTACCCC TAATTTGGT"
    colambanda inclitate da incepta de la colambata de la colambat
  TO A SECTION OF TRANSPORT TARGET AREA CALLED AN TRANSPORTANTAL AND AN AREA SECTION TO A SECTION AS A SECTION 
  CONTROL DESCRIPTION OF A PROPERTY OF A PROPE
  CECCTECT CENTRAL CONTRACTOR CONT
TALE TACEGRACACTECT TALESCATE ANTWASECTICACTICAL SCATE ANTWASECTICAL STATES AND ASSOCIATED TO STATE AND ASSOCIATED TO STATES ASSOCIATED TO STATES AND ASSOCIATED TO STATES AND ASSOCIATED TO STATES ASSOCIATED TO S
CI CCAMMATCÍ CTACCAG
```

FIG. 2-2



F1G. 4

	ptac 5 Promotor Argectational	
348	GINASNI GUGINĜI YGINNGE VAINI EGINAI BIJO SOFP FOAFOTATO TOTOTOTIL YSTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO	181
888	SorbiudiyaletheprobinasploudintheileudinthetbibiybiyHisbinalealeageineileulysbiutheileasnbiudiudialebiutepaipargyel 221 Teabaageageekeeegaabalttaaraeenteenaaraeabababaaleageabaaeageenibaaabaabababaataataabaabeetbeaaabbaatabaafa	122
986	Misprovolmisalogiyprollealeproslyginmetarggiuproargslyseraspilealesiyihrihrserihrlouginglusinilogiyirphetihrasharnpropro 261 Catccactecatecatecatecateatecatecaateagaecaateagegaattateatatatatatatatetateataataataataat	192
106		30
228	Tyrvolasparsphetyrlysth Tatqjagaccggttctatabaac	200
318	AlbleudlyProAleAleAleAleUnetMetThrAlaCyedIn6lyVelGlyGlyPreGlyMistysAleArgVelLeu Stop Stop Dtac 5	

oudresertouphosipaim to: ocimilotarteufrasis di scasartactertissecat

DEAC 5 Promot	ancteimottoutytelulpriloand Eaigeanteijaaasagaciateato obijtaphisopiattousiasiusia asaakiatistaseettitasaakaa	inscretateriletadaptioned inscretatericterdiode tintsiuteletadaptioned	phistyialo koptilau aiselualo Celabarcabariiteestearee Telotyhiityaalilopaaaiseese Araaceeeeraiaceaaaaiseesee	apralaciyyiiletpprosetyriyibiyatprobiyaaphable Propartelatolosaidphalatestalyiyalarafiibase Tittiaobaaaattegeetteetalaabeeaageeqaebaattiet	*Prolinasolysalulostyrprologo *Ainaspargaina?ythrtoisorphod :GCCGAtagagaagaagtatafgtatta	aspirtolloudiogismetasmioup Gaiaeadiatiagaaitee
FIG. 5 ptac 5 Promot standing the control of the c	Sercii cii ai istippogi aaapi eudiaitappet gaalaatafi si sii kii ai iai arot sii ketleut ji siuta lioarasiusia Kasaaceacceaceacatii aaacaecateetaan aacaacaacea ti aacaacaa saa saa saa saa saa saa saa sa	11.0Febble1pcluliofrtysargipplioliology toutalpallabarapeljysorpoprosofilology Aleccedaricalparalisatoraliselabethe translated and the state of the	dialoisi)proalaalatetsesiustusetsestatassastatistassastatassassataassastataassastataassastassass	EpieljargelielphiselamelljesipEzithrGluargelaaliaarpkolaaetytyillötpprosertyrlystyalpargproblyalphalpel Papbadresaadarlacaaatartaaaaattgaargacaaacaagagaattititaabaaaattgeecticcialaagaagaecaagecaagaaattitet Tetebaacaaaaaattoaaaaattgaacagaacaacaagagaattititaabaaaatteecciiccialaagaaagaacaagaaga	trkalaproproblubi uaspradspriblipbiulystrksproberbielysbiebiugs op den jugibiuspytebiuspysbiuden serskalapropro Ionbespropkrobiulybii 1711 Trpbi pbi pbi pbi akarabeskeusorbiudiobi pai akpargbiebi pinstelisospadapadropro Aracccccaccababbebestilbebii 1700cbabbababababtabtebtebtabbebebebebebebebe	arppraserioreinoc Appoleutaitarioargiineijeijeimleulyisiudisiquispaagitariiaopappiartoileudiueiusuusuroprocipty Bacccciceicacatiaacataesbeseractaaagaasciciattagaatagaesbesabataataataataaaaagaattiecenquaaa
Glabs	1000	17 77	20 CA11	1016	Annson Acade	Property of the second

2 2 3

=

=

1768

ipcystysthribelectys 341

Inteltatestraliate 38) Lastacaatecaetrae

polyslyselycystrpats 42 Scammadescrettscash ruginserdrigerablufer 46 rsersiubinibraryhis ti TTEAGAGGAGGAGAGGGA

POLYIGIUPTOPHOATGAIP 301 EAAGGAACCCTTJASAGAT

2

ter

iafroxetfingsrateles 181 iacccaretticaccatta

ISCACATECATACATE
ICCACATECATECATE
PRETTATATATACATE
EATCACATAATCACE

fireptyeluprovimietivypiyyypangoriyaaplovoipiaioteipionietyseidelydiadyydatyfanilotyolologiopiopiones. Tycimmagarcraincmiedagiaidiidogaccancartiagaccaamaaracaamacacaagacagacaagacaatooacaaticaaitaicaabaccattiaaaatc tortystretytystratargrotargotatoristratoarptoltyscialoutentratoaratoristratoarings the sortiovotitoregistreliop Idaabkagdaakialekabekaidakdagdiskkakaraagaistaabaka<u>ett ptac</u> 5 Trpmetipjyrtietentipprakoplytypphyratiotoppoliometlouprotiplyangbertphyratalarakojivetialytetoublipytetouhatypa Genigestipieaetigentetentamateaetastakaecetasmaneetioeeaaaaagaeteeteteaeateataaeaaaettagtoaaattoakthoseeaaete biejiejpaiebijiietiveilyebioloulyelyiteulourgbijydelysijelopfecbivajiidprolouficbibateboodulousekiebiodioukafolo Abatiiaiococosajiaadsiaaacestraidiaadgicciitachecaaceetraacegaatacatacaaacaaacaacaacaacagaaaciocaaaaacacocaa Asaloulouln'statiostyppinta asandaepesitasoppesitasoppesitation of the saloulouln's and saloulous saloulous sa Atchesteacteasiteesiteia talatiteee salous eta seesta salous salous salous salous salous salous salous salous s givlyilityiaislovyigiviityilkebinkebiyyigivojylyisisoriyiiigipesiyendiyandeotyaaistkebibkabololyiyispagor 223 Aaaaataaaacattactababataigtacabaaatgaaaggaaggaaggaagatticaaaaattgegeetgaaateeatataaaatteebiitytyitagtataaaaaabaaaggaga The pripates and the practical tendangs and the languages of the contract of t (ebappiele) japai eferdegevesprotoutopy aappbatgi piprinkai dopenti en saaaandi utbropesi yliekegivedi yrdiaiyrantol Tegateigesigaigeatei itteetiteeettabaagat titagaadetti seesti seestaateataagaataagaataabaatateabiateabiateabia lentegingi firitisti pieteali ilepaginietigepleteliteli (Iqqubi uterbantej tegindeteal) ilotelioterkolapdi plut I egealagoatebaagai ezelelatati egaabatagat egealaanateitagageti itagaaa egaategagatati ilotelaatagaigati itotab tolet protocoeletioet per metapatatatatological contrete mite clonerat per protocopet per per metapatatator des Trestetes ette crates and angenes angenes and consente contrete metapatatate se angenes and angenes and se and strolysfelysmetile 103 atgalaccalantgatag 3110 = :: 500 1001 3 100 = 133

8 ı p16 GAG

/561vLvs	STEGFAGE	
PheAsnC	TTCAACT	
Jallvscv	STTAAGTE	
rotvsThr	GAAAGACC	
okenGlnA	CHACAAA	5.1
3 AsnPheAr	AACTTCAG	hinfi.
ilnAraGlv	AAAGAGGT	10 mn11, 23 hinf1
5' Hete	TATTATGCAAGAGAGGACTTCAGGATCAAGAAGAAGGACGGTTAAGTGTTTCAAGTGTGTAAG	-
	•	

ಅಲ

CTAC KaTE CAAC

E9

₩

ArggluglyHisglnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle AgagaaggicaccaaatgaaggabtgtAccgaaagacaagctaacttcttgggtAagatc TCTCTTCCagtGGTTTACTTCCTGACATGGCTTTCTGTTCGATTGAAGAACCCATTCTAG alut 89 sdul, ban2 hgiA hgiJ11 saot 88 dde1, 76 123

129 bstE2, sau3a, 183

178 bg111 xho2,

alu1,

rsal, 161

148

hph,

131

acc1, 199 ball ofri hael, 184 haeili, 183 rf1,

80

800r11

apy1

204

ProproglugluberpheArgpheglyglugluLysThrThrProserglnLysglnGlu ccacctgaagaagtttcaaggttcgafgaagaagacccccatctcaaaagcaagaa ggtggacttctttcaaagtccaagccactttttttggtggtagggtatttcgttt hph, 270 mbo11, 267 mbo11, 249 243

ProlleAsplysGluleuTyrProleuThrSerleuArgSerleuPheGlyAsnAspPro CCAATCGACAAGGAATTGTACCCATTGACTCTTTGAGATCCTTGTTCGGTAACGATCCC GGTTAGCTGTTCCTTAACAGGTAACTGGAGAAACTCTAGGAACAAGCCATTGCTAGGG 357 sau3a, 340 xho2, 339 307 taq1, 320 raal, 331 mnl1, 361 aval xho1, 303

Berberginop am ' Tcgagccaatgatag Agctcggttactatcagct £9£

9al1 acct hind11 377 taq1

	PYK Promoter Argiciagnalical gracisciacagnanatigicaccachetificacagnatiatigicaccachetifitategagractore	
==	Trplyselualetrtribrleuphocysaleseraspaleargaletyrasptreluvelmisainveltrpaletrimisalecysvelprothraspproainprocla Tegargargargagagagagagagagagagagagagagaga	
> U	VƏLLGUGIYALNVAITHFGIUASHPRASHRETTFÇLYSASHASHRETÇƏLGINGINASPİLƏSIPELENTFPASPĞINSƏFLƏULYSFOÇYSVƏLLYSFOLTFFG 131 Giaticggaatgtgacagaaaattitaacatgiggaaaaataacatggagatgcaggaggatataatcagtttatgggatcaaggctaaagccaigtgtaaaattaaqgcca	=
ں ب	Leucysvelthileuaincysthraipleuglylysalethrainthrasnserseraintplysglugluljelysglyglullelysasncysserpheaintlethrthrserlig 171 Cicicicitactitaaatigcactgattiggggaaggctaataatacaatagtaatiggaagaagaataaaggagaaataaaaactgctgttitgaatatcaccacaaggta	~
a a	argatplytilegialvatalieleupheargasaleuatputtaiprojieaspasaalasetartartaratarataratareleuilentscysasaargserui 2 Aragataaraticagaargaaatecactitticotaacctafotatetartarataratactactactactarcaactataccaactatargiumistarcattetaaca	_
	lietkelhalecysprolysvolserphediuprolleprollemistyrcystrrproaleglyphealolielulyscysashashlysthrpheashglypyselyprocysthr 251 Attacacacacatarabatatatatatatatatatatatatatatatatat	-
	Ainyelsetthyelbingysthrhisölyileargfroileysisetthröinleuleelevasnbiyserleudisbiubiubivelveltieargsgraspasnphethrasnasn 291 Aatotcabcacaatatacatatabaattabbcaatagtbicaactcaactbittaaatbbcabtctagcababbabbingtabattababaatttcacbaacaat	=
	AIƏLYSTHY II OII OVOIGINLOUASNGIUSOYVSIAI ƏII OASNÇSTHYARASNASNASNASNASNASNASNATI OOLI OOLI OOLI OOLI OOLI OOLI Getaaaccataatagtacactgaatgaatetatecaattaactgtacaagacccaacaatacaagaaaqatatetatatagaaccaggaagacatacaagaaga	=
	itoliogiyätpitaatalasismistysatulisseratalogintpäsääsntkilouginitovollysilouarggiuginphogiyäshänlysikrilovoiph Ataataggagatataagaaagcacattataagagacaatagagagaataacactttagaagagatagtaaaaattaagagagttigggaataataaagaatagtett	•
	Ainginsersergiyaiyardikiletdinethisserpheasnaysarggiydiuphephetyraysarntrithebheasnasntritpargleuainhtithagu Aatcaatcotcaggaggaccagaaattgtaatgcagagtttaattgtagagggaatttttotagtgaatagaagagaggtgataataatagagggttaaatgagaggtgaa	-
	Glythlysgiyasnaspthritelleleuprocysargitelysginitellearmettrpgingtuvetblyksatemetyraleproprojjegtygiygtilgeorcysser 451 Gearctaaaggaatgacaatgatgactggatgataaaagaaattataaagatgaggaggaagtaggaagaa	=
••	Seraeniiithelyleuleuleuthargaipliyliythashyeithrashaspihteiweitheargprosiysiysiyaspmeiargaintrargsersiuleutyrlys Tcaatatialassestestattaashasagatostasaatotaastasassessessestetsassessessassassassassassastaattatataaa	=
	Tyrlysvellielysilegiuprodasservelser. Tatadagtaataaaaaattgaatgaattgagaat	

FIG. 7

Nucleotide Positions		FIG. 8-1
Relative to Figure 2	1	MetlleAspLysAlaGlnGluGluHisGluLysTyrHisSerasnTrp AGGXAACAG::::ATGAT:GA:AAGGCACAAGAAGAACATGAGAAATATCACAGTAATTGG TCCXTTGTC::::TACTA:CT:TTCCGTGTTCTTCTTGTACTCTTTATAGTGTCATTAACC
		32 mbo11, 38 nla111,
3820	. 62	ArgalametalaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSer AGAGCCATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGC TCTCGGTACCGATCACTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTCG
•.		66 ncoi, 67 nlaiti, 118 nspBII pvuit, 119 alui,
3880	122	Cysasplyscys61nLeulysG1yG1uA1aMetHisG1yG1nVa1AspCysSerProG1y TGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGA ACACTATTTACAGICGATTTTCCTCTTCGGTACGTACCTGTTCATCTGACATCAGGTCCT
	•	135 alu1, 151 nla111, 152 nsi1 ava3, 155 nla111, 164 acc1, 1 76 apy1 bstX1 ecor11 scrF1,
3940	182	IletroginleuaspCysThrHisLeuGluGlyLysIlelleLeuValAlaValHisVal Atatggcaactagattgtacacttagaaggaaattatcctggtagcagttcatgta Tataccgttgatctaacatgtgtagatcttcctttttaaataggaccatcgtcaagtacat
	•	198 rsal, 205 xbai, 223 apyl ecorii scrfi, 236 mlaiii,
4000	242	AlaserglyTyrlleGlualaGluVallleProalaGluThrGlyGlnGluThralaTyr GCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAGACAGGGCAGGAAACAGCATAT CGGTCACCTATATATCTTCGTCTTCAATAAGGTCGTCTCTGTCCCGTCCTTTGTCGTATA
• •		263 xmn1,
4060	302	PheLeuLeuLysLeuAla6lyArgTrpProValLysThrlleHisThrAspAsnGlySer TTTCTCTTAAAATTA6CA6GAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGC AAAGAGAATTTTAATCGTCCTTCTACCGGTCATTTTT6TTAT6TAT6TCTGTTACCGTCG
		321 mbo11, 326 bali cfri haei, 327 haeill, 357 bbv fnu4hi,
4120	362	AsnPheThrSerThrThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPhe AATTTCACCAGTACTACGGTTAAGGCCGCCTGTTGGTGGGCAGGGATCAAGCAGGAATTT TTAAAGTGGTCATGATGCCAATTCCGGCGGACAACCACCCGTCCCTAGTTCGTCCTTAAA
•		366 hph, 371 scal, 372 rsal, 385 haeill, 386 inu4hi nsbil, 4 05 binl, 406 dpni sau3a,
4180	422	GlylleProtyrAsnProglnSerginglyValValGluSerMetAsnAsnGluLeuLys GGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAAG CCGTAAGGGATGTTAGGGGTTTCAGTTCCTCATCATCTTAGATACTTATTACTTAATTTC
		423 bsm1, 458 hinf1,
4240	482	Lysileile6ly6lnValArqAspGlnAlaGluHisLeuLysThrAlaVal6lnMetAla AAAATTATAGGACAGGTAAGAGATCAGGCTGAACACCTTAAGACAGCAGTACAAATGGCA TTTTAATATCCTGTCCATTCTCTAGTCCGACTTGTGGAATTCTGTCGTCATGTTTACCGT
		503 dpm1 sau3a, 518 afl11, \$30 rsaI,
4300	542	ValphelleHisasnPheLysArqLysGlyGlyIleGlyGlyTyrserAlaGlyGlyArq GTATTCATCCACAATTTTAAAAGAAAAGGGGATTGGGGGATACAGTGCAGGGGAAAGA CATAAGTAGGTGTTAAAATTTTCTTTTC
		\$47 _, fok1, 557 aha111,
4360	602	IleValaspileIleAlaThrAspileGinThrLysGluLeuGinLysGinileThrLys ATAGTAGACATAATAGCAACAGACATACAAACTAAAGAACTACAAAAGCAAATTACAAAA TATCATCTGTATTATCGTTGTCTGTATGTTTGATTCTTGATGTTTTCGTTTAATGTTTT
		60S acci.
4420	662	lleglnasnphearqvaltyrtyrarqaspasnlysaspproleutrolysglyproala attcaaaattitcgggtttattacagggacaacaaagatcccctttggaaaggaccagca taagttttaaaagcccaaataatgtccctgttgtttttaggggaaacctttcctggtcgt
		697 xho2, 698 dpn1 sau3a, 713 asu1 ava2, '

122 AAGCTICTCTGGAAGGGGGGGGGGGGGAGAGAGAGGGGGGGGGG
--

FIG. 8-2